

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/712,768

DATE: 11/29/2000  
TIME: 09:05:50

ENTERED

Input Set : A:\SeqList20511.txt  
Output Set: N:\CRF3\11292000\I712768.raw

4 <110> APPLICANT: P. HOFFMANN-LA ROCHE AG  
6 <120> TITLE OF INVENTION: Cytochrome c oxidase and its genes  
8 <130> FILE REFERENCE: 20511 EP  
C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/712,768  
C--> 11 <141> CURRENT FILING DATE: 2000-11-14  
13 <150> PRIOR APPLICATION NUMBER: 99.122842.0  
14 <151> PRIOR FILING DATE: 1999-11-17  
16 <160> NUMBER OF SEQ ID NOS: 18  
18 <170> SOFTWARE: Patentin Ver. 2.0  
20 <210> SEQ ID NO: 1  
21 <211> LENGTH: 1674  
22 <212> TYPE: DNA  
23 <213> ORGANISM: Gluconobacter oxydans  
25 <220> FEATURE:  
26 <221> NAME/KEY: CDS  
27 <222> LOCATION: (1)..(1674)  
29 <400> SEQUENCE: 1  
31 atg gca gac gcc gcc att cac ggc cat gac cac cat gag aag caa ggc 48  
32 Met Ala Asp Ala Ala Ile His Gly His Asp His His Glu Lys Gln Gly  
33 1 5 10 15  
35 ttc ttc acg cgc tgg ttc atg tcg acc aac cac aaa gac atc ggt ctg 96  
36 Phe Phe Thr Arg Trp Phe Met Ser Thr Asn His Lys Asp Ile Gly Leu  
37 20 25 30  
39 cta tac ctt gta gcg gct ggt gtt ggt ttc att tcc gtc ctg ttc 144  
40 Leu Tyr Leu Val Ala Ala Gly Val Val Gly Phe Ile Ser Val Leu Phe  
41 35 40 45  
43 acc gtc tac atg cgc ctt gag ctg atg gat ccg ggt gtt cag tac atg 192  
44 Thr Val Tyr Met Arg Leu Glu Leu Met Asp Pro Gly Val Gln Tyr Met  
45 50 55 60  
47 tgc ctt gaa ggc gca cgt ctg atc gcg gat gcc tcg cag aca tgt acg 240  
48 Cys Leu Glu Gly Ala Arg Leu Ile Ala Asp Ala Ser Gln Thr Cys Thr  
49 65 70 75 80  
51 gcg aac gga cac ctg tgg aac gtc atg gtt acc tac cat ggt att ctg 288  
52 Ala Asn Gly His Leu Trp Asn Val Met Val Thr Tyr His Gly Ile Leu  
53 85 90 95  
55 atg atg ttc ttt gtg ggt atc ccc gca ttg ttc ggt ggt ttt ggt aac 336  
56 Met Met Phe Phe Val Gly Ile Pro Ala Leu Phe Gly Gly Phe Gly Asn  
57 100 105 110  
59 tat ctg atg ccg ctg caa atc ggc gct ccg gat atg gcc ttc ccg cgt 384  
60 Tyr Leu Met Pro Leu Gln Ile Gly Ala Pro Asp Met Ala Phe Pro Arg  
61 115 120 125  
63 atg aac aac ctg tcg ttc tgg ctg ttc att gcc ggt acc gcg atg ggc 432  
64 Met Asn Asn Leu Ser Phe Trp Leu Phe Ile Ala Gly Thr Ala Met Gly  
65 130 135 140  
67 gtg gct tcg ctg ttc gca ccg ggc ggt gac ggt cag ctg ggt tcg ggc 480  
68 Val Ala Ser Leu Phe Ala Pro Gly Gly Asp Gly Gln Leu Gly Ser Gly  
69 145 150 155 160

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/712,768

DATE: 11/29/2000  
TIME: 09:05:50

Input Set : A:\SeqList20511.txt

Output Set: N:\CRF3\11292000\I712768.raw

```

71 gtt ggt tgg gtt ctg tac ccg ccg ctg tcg acc cgc gaa gct ggc tat 528
72 Val Gly Trp Val Leu Tyr Pro Pro Leu Ser Thr Arg Glu Ala Gly Tyr
73 165 170 175
75 tgg atg gac ctg gcg att ttc gcg gtt cac ttg tcg ggt gcc tcc tcg 576
76 Ser Met Asp Leu Ala Ile Phe Ala Val His Leu Ser Gly Ala Ser Ser
77 180 185 190
79 atc atg ggc gcg atc aac atg atc acg acc ttc ttg aac atg cgc gcc 624
80 Ile Met Gly Ala Ile Asn Met Ile Thr Thr Phe Leu Asn Met Arg Ala
81 195 200 205
83 ccc ggc atg acg ctg cac aaa gtc ccg ttg ttc tcg tgg tcg atc ttt 672
84 Pro Gly Met Thr Leu His Lys Val Pro Leu Phe Ser Trp Ser Ile Phe
85 210 215 220
87 atc acg gct tgg ctg atc ctg ctg gcg ctg ccg gtt ctg gct ggt gca 720
88 Ile Thr Ala Trp Leu Ile Leu Leu Ala Leu Pro Val Leu Ala Gly Ala
89 225 230 235 240
91 atc acc atg ctg ctg acc gac cgt aac ttc ggc acg acc ttc ttc aat 768
92 Ile Thr Met Leu Leu Thr Asp Arg Asn Phe Gly Thr Thr Phe Phe Asn
93 245 250 255
95 cct gct ggc ggc ggt gac ccg att ctg tac caa cac atc ctg tgg ttc 816
96 Pro Ala Gly Gly Asp Pro Ile Leu Tyr Gln His Ile Leu Trp Phe
97 260 265 270
99 ttt ggg cac ccg gaa gtg tac atc atc att ctg ccc ggc ttt ggc atc 864
100 Phe Gly His Pro Glu Val Tyr Ile Ile Ile Leu Pro Gly Phe Gly Ile
101 275 280 285
103 atc agc cat gtc gtg tcg acc ttc tcg aaa aaq ccg gtc ttc ggt tac 912
104 Ile Ser His Val Val Ser Thr Phe Ser Lys Lys Pro Val Phe Gly Tyr
105 290 295 300
107 ctg ccg atg gtc tat gca atg gtg gca atc ggt gtt ctg ggc ttt gtc 960
108 Leu Pro Met Val Tyr Ala Met Val Ala Ile Gly Val Leu Gly Phe Val
109 305 310 315 320
111 gtc tgg gcg cac cac atg tac acc gtt ggt atg tcg ctg acc cag caa 1008
112 Val Trp Ala His His Met Tyr Thr Val Gly Met Ser Leu Thr Gln Gln
113 325 330 335
115 tcc tac ttc atg ctg gcc acc atg gtg atc gcg gtg ccg acc ggc att 1056
116 Ser Tyr Phe Met Leu Ala Thr Met Val Ile Ala Val Pro Thr Gly Ile
117 340 345 350
119 aag atc ttc tcg tgg atc gcc acg atg tgg ggc ggc tcg gtt gag ttc 1104
120 Lys Ile Phe Ser Trp Ile Ala Thr Met Trp Gly Gly Ser Val Glu Phe
121 355 360 365
123 aaa tcg ccg atg ctg tgg gcc ttt gcc ttt atg ttc ctg ttc acc gtg 1152
124 Lys Ser Pro Met Leu Trp Ala Phe Gly Phe Met Phe Leu Phe Thr Val
125 370 375 380
127 ggt ggt gtg acc ggt atc gtg ctg gcc caa gcg ggt ctg gac cgt gca 1200
128 Gly Gly Val Thr Gly Ile Val Leu Ala Gln Ala Gly Leu Asp Arg Ala
129 385 390 395 400
130 tat cac gac acc tat tac gtg gtg gcg cac ttc cat tat gtg atg tcg 1248
131 Tyr His Asp Thr Tyr Tyr Val Val Ala His Phe His Tyr Val Met Ser
132 405 410 415
134 ctg ggt gcg atc ttt gcg atc ttc gcc ggt atc tac ttt tac atg ccg 1296

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/712,768  
 DATE: 11/29/2000  
 TIME: 09:05:50

Input Set : A:\SeqList20511.txt  
 Output Set: N:\CRF3\11292000\I712768.raw

```

135 Leu Gly Ala Ile Phe Ala Ile Phe Ala Gly Ile Tyr Phe Tyr Met Pro
136          420          425          430
138 aag ttc tcg ggc cgc gct ttc ccg gaa tgg gct qca aag ctg cac ttc 1344
139 Lys Phe Ser Gly Arg Ala Phe Pro Glu Trp Ala Ala Lys Leu His Phe
140          435          440          445
142 tgg acc ttc ttc atc ggt gcg aac gtc acg ttc ttc ccg cag cac ttc 1392
143 Trp Ihr Phe Phe Ile Gly Ala Asn Val Ihr Phe Phe Pro Gln His Phe
144          450          455          460
146 ctg gga cgt cag ggt atg ccg cgc cgt tac atc gac tat ccc gaa gcc 1440
147 Leu Gly Arg Gln Gly Met Pro Arg Arg Tyr Ile Asp Tyr Pro Glu Ala
148 465          470          475          480
150 ttc gcg ctg tgg aac aaa gtc tcg tcc tat ggt gcg ttc ctg gcc ttc 1488
151 Phe Ala Leu Lrp Asn Lys Val Ser Ser Tyr Gly Ala Phe Leu Ala Phe
152          485          490          495
154 gcc tcg ttc ctg ttc ttc atc gtg atc ttt gtc tat acg ctg gtt gct 1536
155 Ala Ser Phe Leu Phe Phe Ile Val Ile Phe Val Tyr Thr Leu Val Ala
156          500          505          510
158 ggc cgc cgc gag acc cgt ccg aac ccg tgg ggc gaa ttc gcc gat acg 1584
159 Gly Arg Arg Glu Thr Arg Pro Asn Pro Trp Gly Glu Phe Ala Asp Thr
160          515          520          525
162 ctg gaa tgg acg ctg cca tca ccg cct ccg gcc cac acg ttc gaa acg 1632
163 Leu Glu Trp Thr Leu Pro Ser Pro Pro Ala His Thr Phe Glu Thr
164          530          535          540
166 ctg ccc aag cgc tcg gac tgg gac aag cat ccc tcg cac taa 1674
167 Leu Pro Lys Arg Ser Asp Trp Asp Lys His Pro Ser His
168 545          550          555
171 <210> SEQ ID NO: 2
172 <211> LENGTH: 557
173 <212> TYPE: PR1
174 <213> ORGANISM: Gluconobacter oxydans
176 <400> SEQUENCE: 2
179 Met Ala Asp Ala Ala Ile His Gly His Asp His His Glu Lys Gln Gly
180 1          5          10          15
181 Phe Phe Thr Arg Trp Phe Met Ser Thr Asn His Lys Asp Ile Gly Leu
182          20          25          30
183 Leu Tyr Leu Val Ala Ala Gly Val Val Gly Phe Ile Ser Val Leu Phe
184          35          40          45
185 Thr Val Tyr Met Arg Leu Glu Leu Met Asp Pro Gly Val Gln Tyr Met
186          50          55          60
187 Cys Leu Glu Gly Ala Arg Leu Ile Ala Asp Ala Ser Glu Thr Cys Thr
188 65          70          75          80
189 Ala Asn Gly His Leu Trp Asn Val Met Val Thr Tyr His Gly Ile Leu
190          85          90          95
191 Met Met Phe Phe Val Gly Ile Pro Ala Leu Phe Gly Gly Phe Gly Asn
192          100          105          110
193 Tyr Leu Met Pro Leu Gln Ile Gly Ala Pro Asp Met Ala Phe Pro Arg
194          115          120          125
195 Met Asn Asn Leu Ser Phe Trp Leu Phe Ile Ala Gly Thr Ala Met Gly
196          130          135          140

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/712,768

DATE: 11/29/2000

TIME: 09:05:50

Input Set : A:\SeqList20511.txt

Output Set: N:\CRF3\11292000\I712768.raw

```

197 Val Ala Ser Leu Phe Ala Pro Gly Gly Asp Gly Gln Leu Gly Ser Gly
198 145 150 155 160
199 Val Gly Trp Val Leu Tyr Pro Pro Leu Ser Thr Arg Glu Ala Gly Tyr
200 165 170 175
201 Ser Met Asp Leu Ala Ile Phe Ala Val His Leu Ser Gly Ala Ser Ser
202 180 185 190
203 Ile Met Gly Ala Ile Asn Met Ile Thr Thr Phe Leu Asn Met Arg Ala
204 195 200 205
205 Pro Gly Met Thr Leu His Lys Val Pro Leu Phe Ser Trp Ser Ile Phe
206 210 215 220
207 Ile Thr Ala Trp Leu Ile Leu Leu Ala Leu Pro Val Leu Ala Gly Ala
208 225 230 235 240
209 Ile Thr Met Leu Leu Thr Asp Arg Asn Phe Gly Thr Thr Phe Phe Asn
210 245 250 255
211 Pro Ala Gly Gly Asp Pro Ile Leu Tyr Gln His Ile Leu Trp Phe
212 260 265 270
213 Phe Gly His Pro Glu Val Tyr Ile Ile Ile Leu Pro Gly Phe Gly Ile
214 275 280 285
215 Ile Ser His Val Val Ser Thr Phe Ser Lys Lys Pro Val Phe Gly Tyr
216 290 295 300
217 Leu Pro Met Val Tyr Ala Met Val Ala Ile Gly Val Leu Gly Phe Val
218 305 310 315 320
219 Val Trp Ala His His Met Tyr Thr Val Gly Met Ser Leu Thr Gln Gln
220 325 330 335
221 Ser Tyr Phe Met Leu Ala Thr Met Val Ile Ala Val Pro Thr Gly Ile
222 340 345 350
223 Lys Ile Phe Ser Trp Ile Ala Thr Met Trp Gly Gly Ser Val Glu Phe
224 355 360 365
225 Lys Ser Pro Met Leu Trp Ala Phe Gly Phe Met Phe Leu Phe Thr Val
226 370 375 380
227 Gly Gly Val Thr Gly Ile Val Leu Ala Gln Ala Gly Leu Asp Arg Ala
228 385 390 395 400
229 Tyr His Asp Thr Tyr Tyr Val Val Ala His Phe His Tyr Val Met Ser
230 405 410 415
231 Leu Gly Ala Ile Phe Ala Ile Phe Ala Gly Ile Tyr Phe Tyr Met Pro
232 420 425 430
233 Lys Phe Ser Gly Arg Ala Phe Pro Glu Trp Ala Ala Lys Leu His Phe
234 435 440 445
235 Trp Thr Phe Phe Ile Gly Ala Asn Val Thr Phe Phe Pro Gln His Phe
236 450 455 460
237 Leu Gly Arg Gln Gly Met Pro Arg Arg Tyr Ile Asp Tyr Pro Glu Ala
238 465 470 475 480
239 Phe Ala Leu Trp Asn Lys Val Ser Ser Tyr Gly Ala Phe Leu Ala Phe
240 485 490 495
241 Ala Ser Phe Leu Phe Phe Ile Val Ile Phe Val Tyr Thr Leu Val Ala
242 500 505 510
243 Gly Arg Arg Glu Thr Arg Pro Asn Pro Trp Gly Glu Phe Ala Asp Thr
244 515 520 525
245 Leu Glu Trp Thr Leu Pro Ser Pro Pro Pro Ala His Thr Phe Glu Thr

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/712,768

DATE: 11/29/2000

TIME: 09:05:50

Input Set : A:\SeqList20511.txt

Output Set: N:\CRF3\11292000\I712768.raw

```

246      530      535      540
247 Leu Pro Lys Arg Ser Asp Trp Asp Lys His Pro Ser His
248 545      550      555
251 <210> SEQ ID NO: 3
252 <211> LENGTH: 132
253 <212> TYPE: DNA
254 <213> ORGANISM: Gluconobacter oxydans
256 <220> FEATURE:
257 <221> NAME/KEY: CDS
258 <222> LOCATION: (1)..(132)
260 <400> SEQUENCE: 3
263 ccg ctg gaa atc gtc tgg acg att gtt ccg gtt gtg att ctg gtc ttc 48
264 Pro Leu Glu Ile Val Trp Thr Ile Val Pro Val Val Ile Leu Val Phe
265 1      5      10      15
267 atc ggt gcg ttc tgg ctg ccg gtg ctg ttc aaa cag caa gag ttc ccc 96
268 Ile Gly Ala Phe Ser Leu Pro Val Leu Phe Lys Gln Gln Glu Phe Pro
269      20      25      30
271 gag ggt gac atc gtc atc aac gtc gag ggt cgt agc      132
272 Glu Gly Asp Ile Val Ile Asn Val Glu Gly Arg Ser
273      35      40
276 <210> SEQ ID NO: 4
277 <211> LENGTH: 44
278 <212> TYPE: PRI
279 <213> ORGANISM: Gluconobacter oxydans
281 <400> SEQUENCE: 4
284 Pro Leu Glu Ile Val Trp Thr Ile Val Pro Val Val Ile Leu Val Phe
285 1      5      10      15
286 Ile Gly Ala Phe Ser Leu Pro Val Leu Phe Lys Gln Gln Glu Phe Pro
287      20      25      30
288 Glu Gly Asp Ile Val Ile Asn Val Glu Gly Arg Ser
289      35      40
292 <210> SEQ ID NO: 5
293 <211> LENGTH: 114
294 <212> TYPE: DNA
295 <213> ORGANISM: Gluconobacter oxydans
297 <220> FEATURE:
298 <221> NAME/KEY: CDS
299 <222> LOCATION: (1)..(114)
301 <400> SEQUENCE: 5
304 atc gtc cac gcc gac cgc aag aaa acc gcc att gcc cta gcc att gcc 48
305 Ile Val His Gly Asp Arg Lys Lys Thr Ala Ile Gly Leu Ala Ile Ala
306 1      5      10      15
308 atc gcc att gcc tgg atc ttt acc ctg tgc caa gcc tat gaa tat tat 96
309 Ile Gly Leu Gly Trp Ile Phe Thr Leu Cys Gln Ala Tyr Glu Tyr Tyr
310      20      25      30
312 gaa atc gtc cat acc gaa
313 Glu Ile Val His Thr Glu      114
314      35
317 <210> SEQ ID NO: 6

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/712,768

DATE: 11/29/2000  
TIME: 09:05:51

Input Set : A:\SeqList20511.txt

Output Set: N:\CRF3\11292000\I712768.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date